# The ☺ Joy ☺ of Description Logics

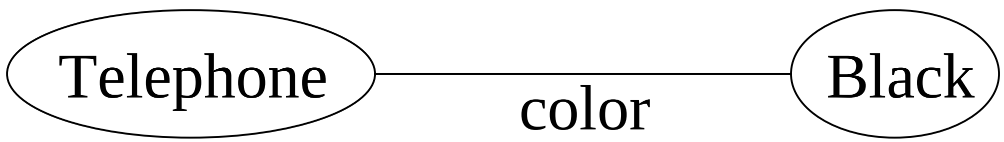
**Scott Sanner**
**(*ssanner@cs.toronto.edu*)**

# Talk Outline

- Motivations and history
- Importance of subsumption and taxonomy
- Semantics for conceptual logics
- Algorithms for subsumption
- Example applications to IR/NLP and Semantic Web
- Frontiers for research

# Motivations and History

- In the beginning (60's & early 70's): knowledge representation (KR) focused primarily on **semantic networks**, e.g.



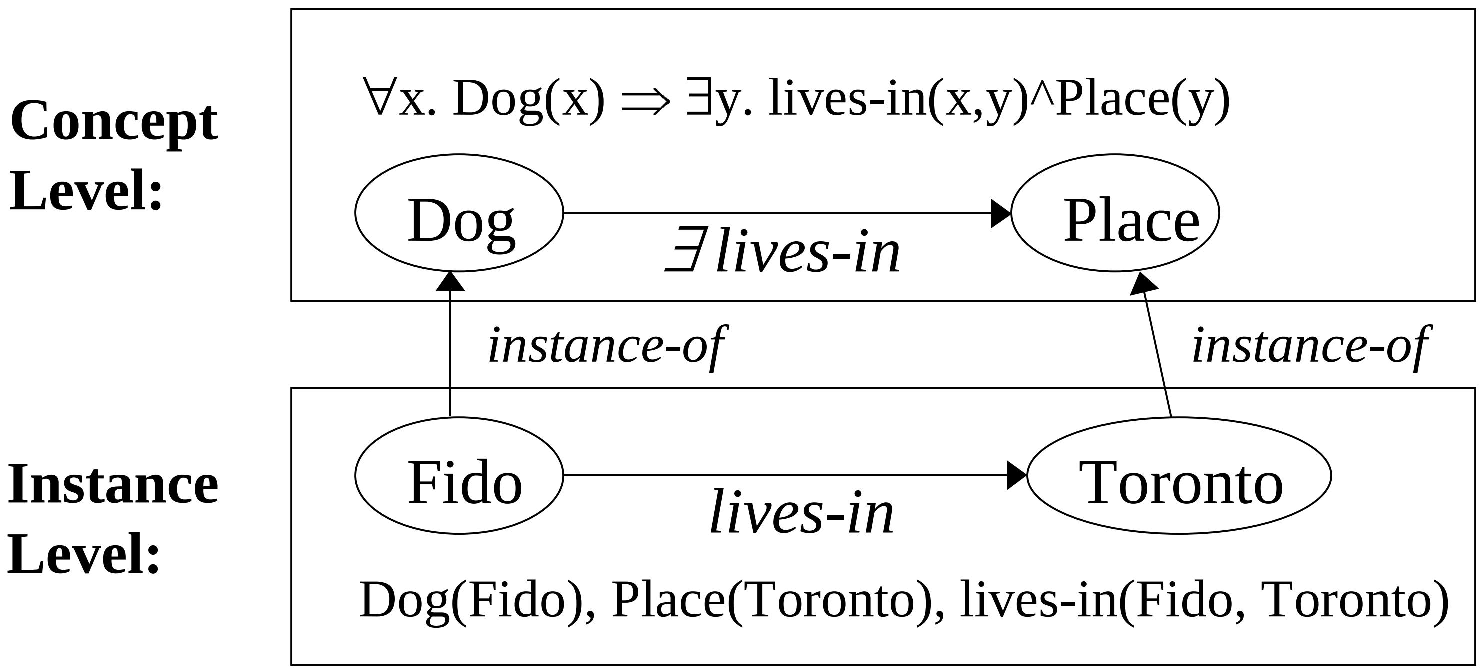

- As pointed out by Woods (1975), what does this semantic net denote?
  - If it's a telephone, it's black?
  - A concept consisting of all black telephones?
  - An instance of a black telephone?

# Motivations and History

- So all was not well in KR, **semantic nets seemed to lack formal semantics**, until…
  - **"What's in a Link" (Woods, 1975):** Provides a *logical foundation* for semantic networks
  - **Structured Inheritance Networks (Brachman, 1977):** Provides further foundations for *structured concepts, subsumption, and automatic taxonomic classification*
  - **Initial KL-ONE Proposal (Woods & Brachman, 1977):** Provides *first conceptual language* based on these ideas

# Motivations and History

- **Ideas from “What’s in a Link”:**
  - Identifies **difference between concept and instance level** (i.e. instances are the extension of concepts)
  - Points out the need for **quantificational import (∀,∃) in concept-level links** (i.e. role restrictions)

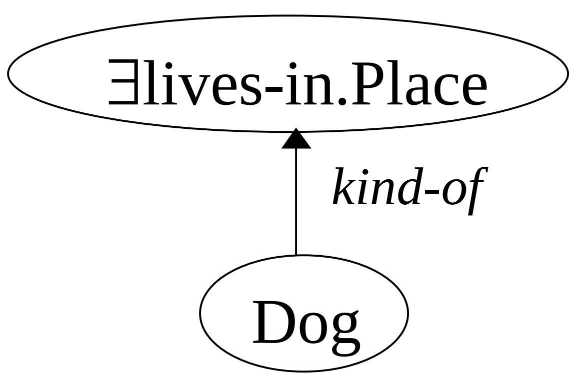


# Motivations and History

- **More ideas from "What's in a Link":**
  - Points out **distinctions between assertional ($\Rightarrow$) and structural links (**$\Leftrightarrow$)

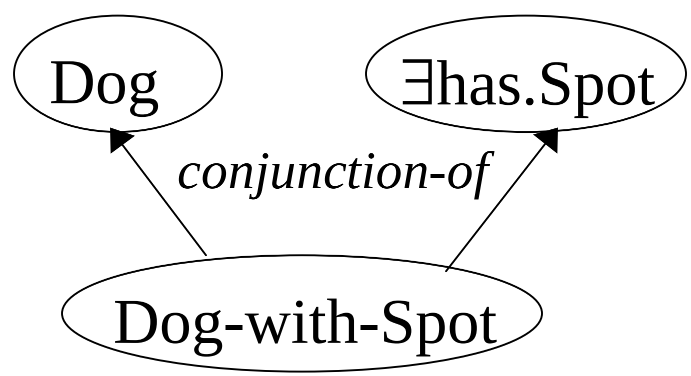
**Assertional:**

$\forall x.\ Dog(x) \Rightarrow$
$\exists y.\ lives\text{-}in(x,y) \wedge Place(y)$




**Structural:**

$\forall x.\ Dog\text{-}with\text{-}Spot(x) \Leftrightarrow$
$Dog(x) \wedge (\exists y.\ has(x,y) \wedge Spot(y))$



# Motivations and History

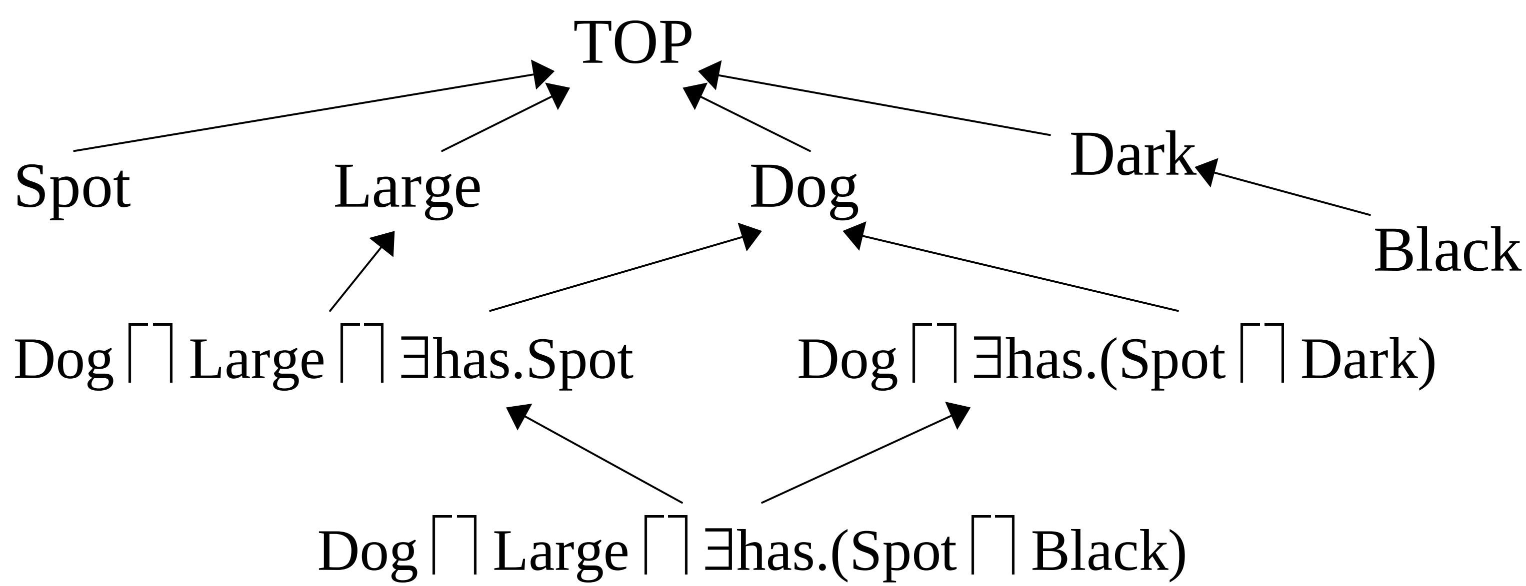
**Aside:** We now have the foundations for a convenient **concept notation** (actually description logic):

| English | FOL | DL |
|---|---|---|
| Dog with a Spot<br>(DWS) | DWS(x) $\Leftrightarrow$<br>Dog(x) ^ ($\exists$y.has(x,y)<br>^ Spot(y)) | DWS $\Leftrightarrow$<br>Dog $\sqcap$ $\exists$has.Spot |
| Large Dog with a Dark Spot<br>(LDWDS) | LDWDS(x) $\Leftrightarrow$<br>(Dog(x) ^ Large(x)) ^<br>($\exists$y.has(x,y)<br>^ (Spot(y) ^ Dark(y)) | LDWDS $\Leftrightarrow$<br>Dog $\sqcap$ Large $\sqcap$<br>$\exists$has.(Spot $\sqcap$ Dark) |

*Note that since variables are not explicit in the DL notation, all relational restrictions are necessarily independent

# Motivations and History

- **Ideas from "Structured Inheritance Networks" and KL-ONE Proposal:**
  - Proposes idea that **<u>structured concepts</u> may subsume each other based on definitional constituents**
  - Furthermore, this **subsumption relation can be used to organize concepts into a taxonomy** (i.e. partial order)



# Motivations and History

- **Subsumption** and the related issue of **taxonomic classification** are perhaps **two of the most beautiful and original ideas in knowledge representation**
  - **Semantics** for such relationships **is criterial** and can be inferred from definition
  - Provides a **mechanism for automatically generating specificity/generality hierarchies**
  - Because real-world information often represented at differing levels of specificity/generality, **leads to many uses in IR, NLP, Information Integration**

# Motivations and History

- **Problem:**
  - Subsumption and taxonomy are important concepts
  - But no formally defined semantics for conceptual logics or algorithms for subsumption computation so far
- **History:**
  - Semantics critical to the definition of algorithms for subsumption
  - Seminal work by Levesque and Brachman (1985) showed that slight alterations in language expressiveness led to extreme changes in computational tractability
- **Answer:**
  - There is no single answer for semantics and subsumption algorithm
  - Virtually all subsequent work has focused on semantics and computational tractability
  - This has led to two major veins of research…

# Motivations and History

- **Two main approaches to conceptual logic semantics and subsumption:**

  **1) Extensional Approach:**
  - Direct subset of FOL model-theoretic semantics
  - *A* subsumes *B* iff there is no model of $B \sqcap \neg A$
  - Basis for most recent description logic research

  **2) Intensional (Structural) Approach:**
  - Assumes an intensional definition of concepts
  - For conjunctively defined concepts:
    *A* subsumes *B* iff every conjunctive constituent of *A* subsumes some conjunctive constituent of *B*
  - Can provide intensional definitions for disjunction, negation, and restriction subsumption as well…

# Summary: A Brief History of Time

**Big Bang**

**70's, Early 1980's: Logical foundations & KL-One**

Ambiguous semantics, undecidable, but fundamental ideas that gave birth to a field

**1985: Levesque & Brachman: Semantics & Tractability**

Formal Model Theoretic Semantics, Tradeoff b/w expressiveness / tractability

**1989: AT&T Research: Classic**

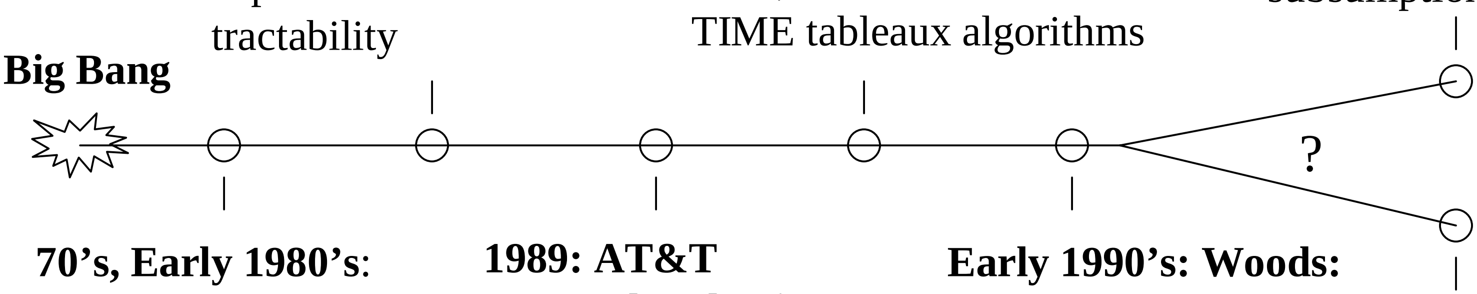
Most expressive, extensionally complete polynomial subsumption language (using structural algorithm)

**Early 1990's: Europe: Need expressiveness and extensional completeness**

Thus, have to use EXP-TIME tableaux algorithms

**Early 1990's: Woods: Need intensional semantics and efficient taxonomy algorithms**

Use structural subsumption (sound w.r.t. extensional semantics) and provide scalable, polynomial tax classification algorithms




**Late 90's, Present: Horrocks & Others: Description Logics, FACT**

Lots of theoretical results and expressive (but complex) sat-based EXP-TIME subsumption algorithms.

**Late 90's, Present: Sun Research: Conceptual Indexing**

Extremely efficient structural algorithms for building large taxonomies, applied to NLP-based web search

# Semantics and Subsumption

- **Extensional Semantics**:

**Note**: Under an interpretation *I*, concepts are just sets of satisfying instances!

| Constructor | Syntax | Semantics |
|---|---|---|
| concept name | $A$ | $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$ |
| top | $\top$ | $\Delta^{\mathcal{I}}$ |
| bottom | $\bot$ | $\emptyset$ |
| conjunction | $C \sqcap D$ | $C^{\mathcal{I}} \cap D^{\mathcal{I}}$ |
| disjunction ($\mathcal{U}$) | $C \sqcup D$ | $C^{\mathcal{I}} \cup D^{\mathcal{I}}$ |
| negation ($\mathcal{C}$) | $\neg C$ | $\Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$ |
| universal | $\forall R.C$ | $\{x \mid \forall y : R^{\mathcal{I}}(x,y) \rightarrow C^{\mathcal{I}}(y)\}$ |
| existential ($\mathcal{E}$) | $\exists R.C$ | $\{x \mid \exists y : R^{\mathcal{I}}(x,y) \wedge C^{\mathcal{I}}(y)\}$ |
| cardinality ($\mathcal{N}$) | $\geqslant n\ R$ | $\{x \mid \sharp\{y \mid R^{\mathcal{I}}(x,y)\} \geq n\}$ |
| | $\leqslant n\ R$ | $\{x \mid \sharp\{y \mid R^{\mathcal{I}}(x,y)\} \leq n\}$ |
| qual. cardinality ($\mathcal{Q}$) | $\geqslant n R.C$ | $\{x \mid \sharp\{y \mid R^{\mathcal{I}}(x,y) \wedge C^{\mathcal{I}}(y)\} \geq n\}$ |
| | $\leqslant n R.C$ | $\{x \mid \sharp\{y \mid R^{\mathcal{I}}(x,y) \wedge C^{\mathcal{I}}(y)\} \leq n\}$ |

Source: http://www.cs.man.ac.uk/~franconi/dl/course/slides/prop-DL/propositional-dl.pdf

# Semantics and Subsumption

- **Sample tableaux for extensional subsumption (i.e. unsatisifiability check)**:
  - Does ¬∃CHILD.¬Male subsume ∀CHILD.Male?
  - I.e., Does ∀CHILD.Male ⊓ ∃CHILD.¬Male have no model?

**Tableaux Proof:**

| | |
|---|---|
| $x : \big((\forall \texttt{CHILD.Male}) \sqcap (\exists \texttt{CHILD.}\neg\texttt{Male})\big)$ | |
| $x : (\forall \texttt{CHILD.Male})$ | ⊓-*rule* |
| $x : (\exists \texttt{CHILD.}\neg\texttt{Male})$ | " |
| $x \; \texttt{CHILD} \; y$ | ∃-*rule* |
| $y : \neg\texttt{Male}$ | " |
| $y : \texttt{Male}$ | ∀-*rule* |
| $\langle CLASH \rangle$ | |

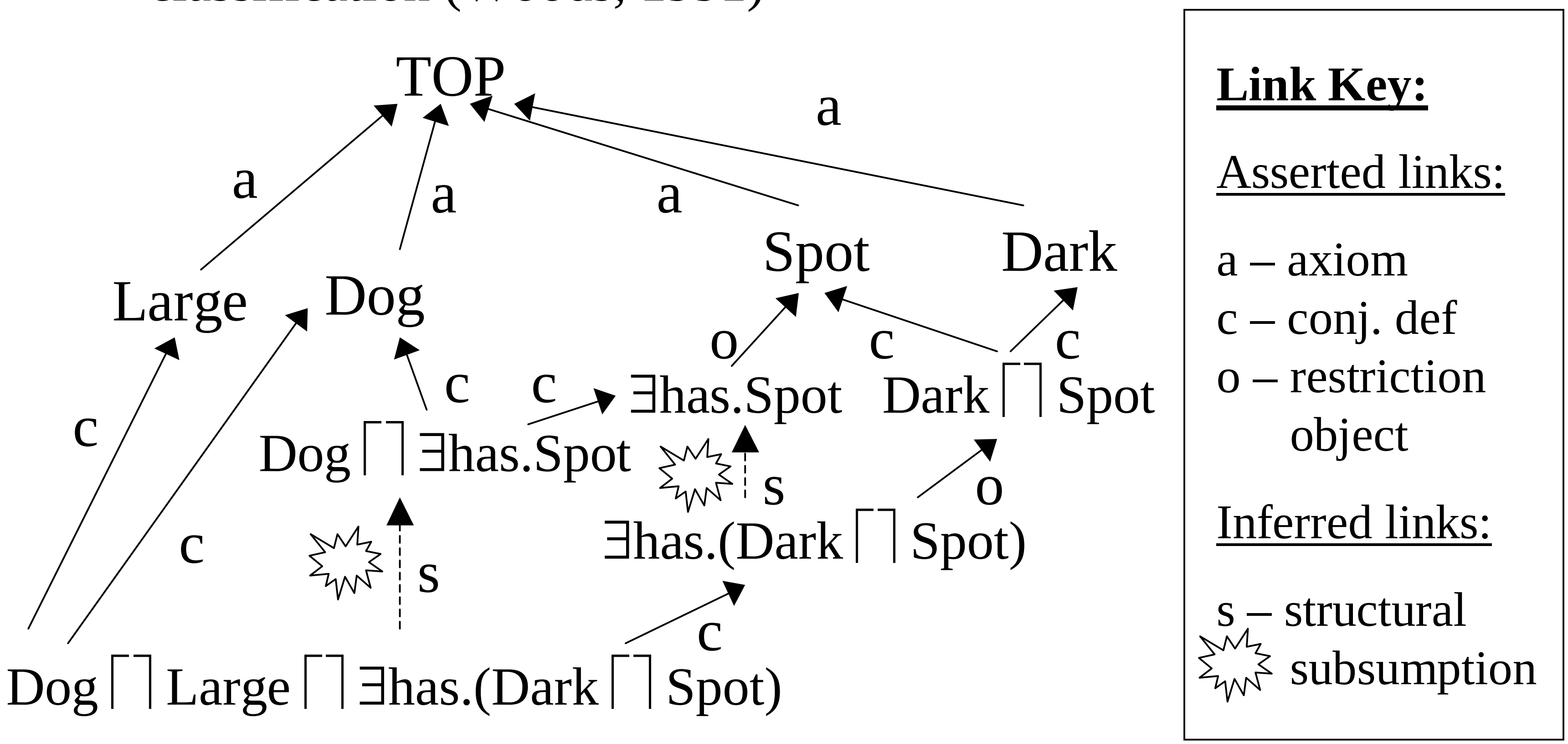
Source: http://www.cs.man.ac.uk/~franconi/dl/course/slides/prop-DL/propositional-dl.pdf

# Semantics and Subsumption

- **Intensional Approach**:
  - Avoid model theoretic semantics, **reason directly using recursive application of subsumption rule to normalized concept structure:**
    - **Conjunctive Subsumption**: A *subsumes* B iff every conjunctive constituent of A *subsumes* some conjunctive constituent of B
    - **Existential Restriction Subsumption**: ∃R.C *subsumes* ∃R'.C' iff R *subsumes* R' and C subsumes C'
    - **Definitions for other restrictions** (Woods, 1991), **disjunction** (Sanner, 2003); various ways to handle **complement**
  - Note the **recursive use of the term *subsumes***
    - **Base cases**: axiomatic and definitional (e.g. conj.) subsumptions
    - **Recursive cases**: structural subsumptions

# Semantics and Subsumption

- **Sample application of intensional (structural) subsumption algorithm:**
  - General rule covers single subsumption test but can be incorporated into an efficient algorithm for taxonomic classification (Woods, 1991)



# Semantics and Subsumption

- **Intuition:** Structural subsumption akin to a constrained second-order search for first-order logic proofs:
  – Previous structural inference can be recast as FOL proof:

| | |
|---|---|
| [1] Large-dog-that-has-a-dark-spot(x) ⇒ Dog(x) | Skolemized INF Def |
| [2] Large-dog-that-has-a-dark-spot(x) ⇒ has(x,f(x)) | Skolemized INF Def |
| [3] Large-dog-that-has-a-dark-spot(x) ⇒ Spot(f(x)) | Skolemized INF Def |
| [4] Dog(x)^(has(x,y)^Spot(y)) ⇒ Dog-that-has-a-spot(x) | Skolemized INF Def |
| [5.1] Large-dog-that-has-a-dark-spot(x) | Assumption |
| [5.2] Dog(x) | MP [5.1],[1] |
| [5.3] has(x,f(x)) | MP [5.1],[2] |
| [5.4] Spot(f(x)) | MP [5.1],[3] |
| [5.5] Dog-that-has-a-spot(x) | MP [4],[5.2-5.4] |
| [6] Large-dog-that-has-a-dark-spot(x) ⇒ Dog-that-has-a-spot(x) | Deduction Theorem [5.1],[5.5] |

Q.E.D.

# Semantics and Subsumption

- **How do extensional and intensional (i.e. structural) approaches/algorithms compare?**

**Intensional (Structural)**

- Sound but incomplete for very expressive languages, however incompleteness often benign
- Efficient (poly-time) taxonomy building
- Expressiveness no longer an issue with recent research
- Captures important subsumptions

Ex: Nova (Sun)
JTP (KSL)

- Can be extensionally sound & complete using structural algorithm
- ⇒ Polynomial inference, but limited expressiveness
- Ex: Classic (AT&T)

**Extensional**

- Complete inference for expressive languages
- Intractable (EXP-TIME) subsumption & taxonomy building
- Ex: FaCT (Horrocks)

# Applications to IR and NLP

- **Nova Conceptual Search Engine** (Sun Microsystems Research Labs)
  - Converts phrases parsed from natural language documents into conceptual logic descriptions
  - Organizes these descriptions in a taxonomy using *highly optimized* structural classification algorithms
  - Answers queries by returning all concepts equivalent to or more specific than query, e.g.
    - Given query: ‘wash an automobile’, will return a document containing the phrase: ‘washing a car with a hose’
    - Given query: ‘run a farm’, will return a document containing the phrase: ‘operating a large dairy’

# Applications to IR and NLP

- **Example Nova taxonomy generated from Sun catalog:**

```
(ADD MEMORY)
  |-k- (ADDITIONAL MEMORY)
  | |-k- (ADDITIONAL A MEMORY)
  | |-k- (ADDITIONAL G MEMORY)
  | |-k- (ADDITIONAL K MEMORY)
  | |-k- (ADDITIONAL STORAGE)
  | |   |-k- (ADDITIONAL DISK)
  | |     |-k- (ADDITIONAL DISKS)
  | |     | |-k- (TWO ADDITIONAL HARD DISKS)
  | |     |
  | |     |-k- (ADDITIONAL MULTI-DISK)
  | |       |-k- (ADDITIONAL 4.2-GB MULTI-DISK)
  | |       |
  | |       |-k- (ADDITIONAL SMCC MULTI-DISK)
  | |
  | |-k- (PURCHASE ADDITIONAL MEMORY)
  |
  |-k- (ECONOMICALLY ADDING LOCAL STORAGE)
  |-k- (SOLDERED-IN MEMORY)
```

Source: http://research.sun.com/knowledge/examples.html

# Applications to IR and NLP

- **For More Information…**
  - Core algorithms are proprietary but publicly available research information can be found at:
    - Nova Project Home Page:
      - http://research.sun.com/knowledge/index.html
    - Sub Labs Tech Report:
      - William A. Woods (1997). "Conceptual Indexing: A Better Way to Organize Knowledge." SMLI-TR-97-61.

# Applications to the Semantic Web

- **JTP DAML+OIL Reasoner** (Knowledge Systems Lab, Stanford University)
  - **Taxonomic reasoning** extremely useful for **determining relationships between concepts in distributed kb's**
  - **Problem: Have a general purpose theorem prover** in the form of JTP (Java Theorem Prover)
    - Capable of loading DAML+OIL knowledge bases from Semantic Web
    - However, inference of a taxonomy by theorem proving is too inefficient
  - **Solution: Build in a special purpose reasoner for DAML+OIL taxonomic classification**
    - Use structural subsumption techniques augmented with rules for for expressive languages (e.g. since disjunction important)

# Applications to the Semantic Web

- **Example taxonomy built by JTP special purpose reasoner for a DAML+OIL kb:**

```
            |-  TOP
        |- http://.../dogs.daml#::Thing
        |   |-  TOP  [*]
        |- http://.../dogs.daml#::AnimalOrDarkFur
    |- http://.../dogs.daml#::Animal
    |       |- http://.../dogs.daml#::AnimalOrDarkFur [*]
    |   |- http://.../dogs.daml#::DogOrBrownFurOrBlackFur
    |- http://.../dogs.daml#::DogOrBrownFur
    |   |- ...
    |- http://.../dogs.daml#::DogOrBrownFurOrBlackFur [*]
|- http://.../dogs.daml#::Dog
|   |- [_GEN_ <_MOD_>]
|- [_EXISTS_ <http://.../dogs.daml#::mod,
|              http://.../dogs.daml#::Big,1>]
|   |- http://.../dogs.daml#::DogOrBrownFur [*]
|- http://.../dogs.daml#::BigDogOrBrownFur
http://.../dogs.daml#::BigDog
|- http://.../dogs.daml#::BigDogWithDarkFur
    |- http://.../dogs.daml#::BigDogWithBrownFur
        |-  BOTTOM
```

Source: http://www.cs.toronto.edu/~ssanner/papers/ksl0301.pdf

# Applications to the Semantic Web

- **For More Information…**
  - JTP and the DAML+OIL special purpose reasoner are freely available, see the following web page and extensive technical report:
    - JTP Project Home Page:
      - http://www.ksl.stanford.edu/software/JTP/
    - KSL Tech Report:
      - Scott P. Sanner (2003). "Towards Practical Taxonomic Classification for Description Logics on the Semantic Web." KSL-03-01.

# Frontiers of Description Logics

- **Extensional / Description Logics**
  - Can the average case for subsumption be competitive with structural approaches (depends on algorithms /distributions of concept constructors used)?
  - Can taxonomic classification be performed more efficiently by reusing subsumption information?
- **Intensional / Structural Logics**
  - How to characterize incomplete subsumption?
  - Are we sure these missed subsumptions don't matter? (e.g. is inconsistency useful for some applications?)
  - Further research in efficient subsumption, algorithms and data structures (disk-based) for scalable tax. construction

# References


- **Ronald J. Brachman (1977). "A Structural Paradigm for Representing Knowledge".** Ph.D. Thesis, Harvard University. Note: Structured inheritance networks proposed here.
- **Hector J. Levesque and Ronald J. Brachman (1985). "A Fundamental Tradeoff in Knowledge Representation and Reasoning".** In R. J. Brachman and H. J. Levesque, editors, *Readings in Knowledge Representation*, pages 41-70. Morgan Kaufmann, Inc. 1985.
- **Scott P. Sanner (2003). "Towards Practical Taxonomic Classification for Description Logics on the Semantic Web".** Stanford University Knowledge Systems Lab Techical Report: KSL-03-01.
- **William A. Woods (1975). "What's in a Link: Foundations for Semantic Networks".** In R. J. Brachman and H. J. Levesque, editors, *Readings in Knowledge Representation*, pages 41-70. Morgan Kaufmann, Inc. 1985.
- **William A. Woods and Ronald J. Brachman (1977). "Research in Natural Language Understanding, Quartly Technical Progress Report No. 1".** Bolt, Beranek, and Newman Techical Report: 3742. Cambridge, MA. Note: Initial KL-ONE proposal.
- **William A. Woods (1991). "Understanding Subsumption and Taxonomy: A framework for progress".** In John Sowa (editor) *Principles of Semantic Networks: Explorations in the Representation of Knowledge*. San Mateo, US.
- **William A. Woods (1997). "Conceptual Indexing: A Better Way to Organize Knowledge".** Sun Microsystems Research Labs Technical Report: SMLI-TR-97-61.